Precoder Design with Limited Feedback and Backhauling for Joint Transmission

Tilak Rajesh Lakshmana, Antti Tölli, *Senior Member, IEEE,*
Rahul Devassy, *Student Member, IEEE,*
and Tommy Svensson, *Senior Member, IEEE*

Abstract

A centralized coordinated multipoint downlink joint transmission in a frequency division duplex system requires channel state information (CSI) to be fed back from the cell-edge users to their serving BS, and aggregated at the central coordination node for precoding, so that interference can be mitigated. The control signals comprising of CSI and the precoding weights can easily overwhelm the backhaul resources. Relative thresholding has been proposed to alleviate the burden; however, this is at the cost of reduction in throughput. In this paper, we propose utilizing the long term channel statistics comprising of pathloss and shadow fading in the precoder design to model the statistical interference for the unknown CSI. In this regard, a successive second order cone programming (SSOCP) based precoder for maximizing the weighted sum rate is proposed. The accuracy of the solution obtained is bounded with the branch and bound technique. An alternative optimization framework via weighted mean square error minimization is also derived. Both these approaches provide an efficient solution close to the optimal, and also achieve efficient backhauling, in a sense that the precoding weights are generated only for the active links. For comparison, a stochastic approach based on particle swarm optimization is also considered.

Index Terms

branch and bound, CoMP, limited backhauling/feedback, MSE, precoding, SSOCP, weighted sum rate maximization

Tilak Rajesh Lakshmana, Rahul Devassy, and Tommy Svensson are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden, e-mail: tilak@chalmers.se, devassy@chalmers.se, tommy.svensson@chalmers.se.

Antti Tölli is with the Centre for Wireless Communications, University of Oulu, Oulu, Finland, e-mail: antti.tolli@ee.oulu.fi
.

I. Introduction

In cellular coordinated multipoint (CoMP) transmission systems, the channel state information (CSI) present at the transmitter plays an important role in harnessing the gains for joint transmission CoMP. In the downlink, a group of base stations (BSs) coordinate to coherently serve a group of users being prone to interference [1]-[3]. To mitigate interference in a frequency division duplex (FDD) system, the users need to estimate the CSI based on the downlink pilots from the BSs, and then feed it back to its cooperating BSs (typically to the serving BS). In a centralized architecture, the BSs forward the CSI to a central coordination node, where the CSI from various users are accumulated to form the aggregated channel matrix which is used to design a precoder for mitigating interference. In a decentralized architecture, the users need to share the CSI between the cooperating BSs to form the precoding weights. Sharing of CSI and the precoding weights between the BSs and the central coordination node occurs over the backhaul. This is typically a microwave or an optical fiber link.

A. Previous work

In an FDD system, the overhead of feeding back the CSI of all the cooperating BSs from the user could easily overwhelm the wireless radio interface and the backhaul resources, especially in a centralized architecture. In this regard, absolute and relative thresholding [4] was proposed to limit the CSI feedback. In particular, relative thresholding is a process in which the users only feedback those links that fall within a threshold, say 5 dB, relative to its strongest BS. This results in limited CSI being available for the precoder design. A user centric clustering is performed in [5], which is similar to the relative thresholding performed in our work. However, the aim is finding the optimal tradeoff between total transmit power and sum backhaul capacity via power minimization.

In [6], [7], linear precoding is considered, as it provides a good tradeoff between complexity and performance. With limited CSI, in [6] a linear zero forcing (ZF) precoder with suboptimal power allocation [8] is considered to achieve the backhaul signaling load reduction based on a physical (PHY) layer precoding and a medium access control (MAC) layer scheduling approach. Apart from the PHY and MAC layer approaches, a predefined constrained backhaul infrastructure can be included in the precoder design as in [9]. The ZF approach requires a well conditioned aggregated channel matrix at the central coordination node for channel inversion, which cannot be

guaranteed with limited CSI. This poses constraints on how the users are selected/scheduled, and it makes it harder to achieve efficient backhauling. In this paper, the term efficient backhauling is used to denote the case where the number of precoding weights generated for the active links is equal to the number of CSI coefficients correspondingly available for the active links at the central coordination node. Note that for example with the ZF approach, it is possible to generate the non-zero precoding weights for non-cooperating BSs and require them to be nulled to achieve efficient backhauling. In a centralized architecture, if the central coordination node decides the routing of user data then this will be based on the precoding weights being generated for only the active links, instead of making all the user data to be available at all the cooperating BSs. Such an approach requires efficient backhauling. Another approach to minimize the backhaul user data transfer is to jointly design the precoder and simultaneously minimize the user data transfer in the backhaul based on the quality of service [10]. In [11], a stochastic precoder based on particle swarm optimization (PSO) is designed to achieve efficient backhauling while taking the limited feedback and limited backhaul into account. However, with suboptimal power allocation and with increase in the problem size, the complexity of the algorithm increases as pointed out in [11, sec. 3.4].

Weighted sum rate maximization is a difficult non-convex problem [12], [13]. In this regard, different centralized successive convex approximation (SCA) methods are proposed in [14]-[16]. In [15], a low complexity approximation with faster convergence rate is proposed for a downlink multicell multiple input single output (MISO) system. In [16], a different approximation is used for robust precoding with uncertainty in CSI at the transmitter. The precoders can also be designed via the mean square error (MSE) approach. In [12], [17], it was shown that minimizing the weighted sum mean square error (MSE) is equivalent to the weighted sum rate maximization, where the precoder, receive weight and the receiver MSE weights are alternately optimized. Whenever a central coordination node is not available, then [12], [18] can be used to implement the precoder in a decentralized fashion. Signaling strategies are considered in [18], and also under imperfect channel conditions [19] extending the result from [17]. Also, in [20], a generalized mean square error criterion is used to arrive at a robust linear precoding solution that can handle backhaul constraints with CSI uncertainty. Similar to weighted sum rate maximization, a cross layer queue deviation minimization is considered in [14] where the queue states act as weights for the sum rate maximization, with a different approximation of the signal to interference plus

noise (SINR) constraint.

B. Contributions

In this work, we focus on the design of the precoder in a centralized FDD system with the objective of maximizing the weighted sum rate of the users with perfect but limited CSI feedback, and also under limited backhauling. We use the algorithms developed for the full CSI case, but now we incorporate the limited CSI and the statistical model of interference. In this regard, we propose a conservative precoder design for any combination of user centric clustering with per-antenna power constraint, where the long term channel statistics is incorporated into the optimization problem for the missing links. Here, we model the statistical interference for the unknown CSI as the long term channel statistics in the interference terms for the user. The model is pessimistic in nature, as the Cauchy-Schwarz inequality is applied on the unknown parts that were separated from the total interference. The long term channel statistics is also used for making the routing decisions for the user data in the backhaul as noted in [21].

In our work, we effectively solve the problem of designing a PHY layer precoder with limited information based on the approach in [14]-[16], where we extend the SCA framework to cope with incomplete CSI at the transmitter. Here we consider joint transmission CoMP while [14]-[16] focused on coordinated beamforming. In this regard, we incorporate the pessimistic interference model based on the long term channel statistics (pathloss and shadow fading) into the problem formulation. Alternatively we also include this statistical interference model in the minimization of the weighted MSE [17]-[18]. We also use the long term channel statistics to determine the CSI feedback threshold. Our proposed pessimistic statistical interference modeling is different compared to the previous work [4], [6], [11], [22] where the unavailable CSI are modeled as zeros. The availability of the long term statistics at the coordination node is a valid assumption, as they are available in the existing cellular standards, where the users feedback the received signal strength, more popularly referred to as the received signal strength indicator (RSSI). For example, this is required during handover procedures. In our setup, we consider relative thresholding, a variant of [4], based on this average signal strength at the user.

The main contributions of this work are listed as follows:

- The long term channel statistics (pathloss and shadow fading), based on relative thresholding, are modeled in the precoder design as part of the pessimistic statistical interference in the

SINR ratio.

- We efficiently solve the precoder design problem with the limited feedback and limited backhauling, using a successive second order cone programming (SSOCP). Also, we solve for the case when the long term channel statistics is considered as part of the SOC constraint, instead of neither treating them as zeros nor naively replacing the zeros with this side information. We also achieve efficient backhauling, where the precoding weights are generated only for those links whose CSI was reported.
- As an alternative to the SSOCP, we reformulate the problem via weighted MSE criterion similar to [12], [17], with the use of the proposed long term channel statistics in the variance of the received signal. The results show that it achieves the same weighted sum rate as that of the proposed SSOCP on average. The MSE reformulation requires a higher number of iterations than SSOCP to converge but each sub-problem is simple to solve.
- We characterize the performance of the proposed precoder design using numerical bounds with a variant of the branch and bound technique [13]. The proposed iterative SSOCP algorithm is very close to the optimal provided by the branch and bound method.
- We numerically compare the performance of the proposed iterative algorithm to an existing stochastic algorithm under limited feedback and limited backhauling. In particular we consider PSO, as the overhead of book keeping of variables is far simpler compared to other stochastic algorithms such as ant colony optimization or evolutionary algorithms. It was found that the performance of the PSO is inferior, especially when the problem size is increased.

The paper is organized as follows: the system model is introduced in Section II. In Section III, the precoder design based on SSOCP and MSE are derived. A brief description of PSO is presented, and this section concludes with the branch and bound technique used to bound the performance of SSOCP. Using the derived precoders, the simulation results are presented in Section IV, in terms of the effect of threshold, cell-edge signal to noise ratio (SNR), BS antennas and the SSOCP bounds. Finally Section V concludes the contribution of the paper.

Notation: A scalar variable is denoted as x while X denotes a scalar constant. A vector and a matrix are denoted as $\mathbf{x}$ and $\mathbf{X}$, respectively. A set is denoted in calligraphic font as $\mathcal{X}$ and the cardinality of the set is $|\mathcal{X}|$. The elements of set $\mathcal{X}$ not in set $\mathcal{Y}$ is denoted as $\mathcal{X}\backslash\mathcal{Y}$. The absolute value of $x\in \mathbb{C}$ is denoted as $|x|$ while the $p-$norm of a vector is denoted as $||\cdot||_p$.

The transpose and conjugate transpose of a vector $\mathbf{x}$ is denoted as $\mathbf{x}^T$ and $\mathbf{x}^H$, respectively. The expectation operation on the random variable X is denoted as $E_X[X]$.

II. System Model

Consider a homogenous network cluster consisting of $|\mathcal{B}|$ BSs, each with N_T antennas. The BSs are coordinated to serve $|\mathcal{U}|$ single antenna cell-edge users. The signal received by the uth user is y_u, and it consists of the desired signal and intracluster interference

$$y_u = \sum_{b\in\mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u}x_u + \sum_{i\neq u}\sum_{b\in\mathcal{B}_i} \mathbf{h}_{b,u}\mathbf{w}_{b,i}x_i + n_u, \tag{1}$$

where $\mathcal{B}_u$ is the set of BSs from which the uth user is served. In this model, the intercluster interference is considered to be negligible for the cell-edge users located at the cluster center, and therefore it is not accounted in (1). The channel experienced by the uth user from bth BS with N_T antennas is $\mathbf{h}_{b,u} \in \mathbb{C}^{1\times N_\mathrm{T}}$. The precoding weight for the uth user with normalized data x_u from the bth BS with N_T antennas is $\mathbf{w}_{b,u} \in \mathbb{C}^{N_\mathrm{T}\times 1}$, such that $\mathbf{w}_{b,u} = [w_{b,u}^{(1)}, w_{b,u}^{(2)}, \ldots, w_{b,u}^{(k)}, \ldots, w_{b,u}^{(N_\mathrm{T})}]^T$ where $w_{b,u}^{(k)}$ is the precoding weight on the kth antenna of the bth BS for the uth user, and n_u is the receiver noise at uth user with power N_0.

To incorporate the long term channel statistics, let us first consider the SINR evaluated at the central coordination node for the uth user as

$$\begin{aligned}
\tilde{\gamma}_u &= \frac{\left|\sum\limits_{b\in\mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2}{\sum\limits_{i\neq u}\left|\sum\limits_{b\in\mathcal{B}_i} \mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2 + N_0} \\
&= \frac{\left|\sum\limits_{b\in\mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2}{\sum\limits_{i\neq u}\left\{\left|\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,i} + \sum\limits_{b\in\mathcal{B}_i\setminus\mathcal{B}_u} \overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2\right\} + N_0}
\end{aligned} \tag{2}$$

where the interference terms in the denominator of (2) are split based on relative thresholding, i.e., the set $\mathcal{B}_i \cap \mathcal{B}_u$ denotes the set of BSs that are involved in serving both the uth and the ith user, as the CSI $\mathbf{h}_{b,u}$ falls within the relative threshold window. However, those links that fall outside this threshold constitute the term $\overline{\mathbf{h}}_{b,u}$ where $\mathcal{B}_i\setminus\mathcal{B}_u$ is the set of BSs serving the ith user but not the uth user. The given set $\mathcal{B}_u$ is defined by the relative thresholding algorithm based on the long term channel statistics as summarized in Algorithm 1. To achieve the condition of

efficient backhauling, the precoding weights are generated only for those links for which the users have fed back the CSI.

Algorithm 1 Relative thresholding performed at the user based on the long term channel statistics (pathloss and shadow fading)

1: Set the feedback threshold, $T\,(= 3\,\text{dB}$, for example$)$
2: **for** $\forall u \in \mathcal{U}$ **do**
3: Perform channel measurements of the BSs, $\mathcal{B}$
4: $c = \max_{b\in\mathcal{B}} \left(E\left[||\mathbf{h}_{b,u}||_2^2\right]\right)$
5: **for** $\forall b \in \mathcal{B}$ **do**
6: **if** $\left(c_{\text{dB}} - \left[E\left[||\mathbf{h}_{b,u}||_2^2\right]\right]_{\text{dB}}\right) \leq T$ **then**
7: Include b in the set $\mathcal{B}_u$
8: **end if**
9: **end for**
10: The uth user feeds back the CSI of the set of BSs in $\mathcal{B}_u$
11: **end for**

Now we define a new SINR, $\overline{\gamma}_u$ in (3a), where we replace the unknown channel coefficients with an expectation as in (3a)-(3e). We obtain (3b) by expanding the terms as $|a+b|^2 = |a|^2 + |b|^2 + ab^H + ba^H = |a|^2 + |b|^2 + 2\Re\left\{ab^H\right\}$and taking the expectation inside. Here the hermitian operator is degenerated to a scalar case. Here we focus on $E_{\overline{h}}\left[\overline{\mathbf{h}}_{b,u}\right]$, where $\overline{\mathbf{h}}_{b,u}$ consists of the three random variables, the pathloss, l, the shadow fading, $s_{b,u} \sim \ln\mathcal{N}(0, \sigma^2_{\text{SF}})$, and the small scale fading on the kth antenna is $f^{(k)}_{b,u} \sim \mathcal{CN}\left(0,1\right)$ and $\mathbf{f}_{b,u} \in \mathbb{C}^{1\times N_{\text{T}}} = \left[f^{(1)}_{b,u}, f^{(2)}_{b,u}, \cdots, f^{(N_{\text{T}})}_{b,u}\right]$. The large scale fading and the small scale fading are independent random variables, and $E_f\left[f_{b,u}\right] = 0$, therefore we have $E_{\overline{h}}\left[\overline{\mathbf{h}}_{b,u}\right] = E_{l,s,f}\left[l_{b,u}s_{b,u}\mathbf{f}_{b,u}\right] = E_{l,s}\left[l_{b,u}s_{b,u}\right]E_r\left[\mathbf{f}_{b,u}\right] = \mathbf{0}_{N_{\text{T}}}$. The inequality in (3d) is obtained using the Cauchy-Schwarz inequality $\left|\sum_{j=1}^{N} a_j b_j^*\right|^2 \leq \sum_{j=1}^{N} |a_j|^2 \sum_{j=1}^{N} |b_j|^2$, where $a_j = \overline{\mathbf{h}}_{j,u}\mathbf{w}_{j,i}$ and $b_j = 1, \forall j$. Finally, we obtain (3e) as follows, $E_{\overline{h}}\left|\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2 = E_{\overline{h}}\left[\mathbf{w}^H_{b,i}\overline{\mathbf{h}}^H_{b,u}\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right] = \mathbf{w}^H_{b,i}E_{\overline{h}}\left[\overline{\mathbf{h}}^H_{b,u}\overline{\mathbf{h}}_{b,u}\right]\mathbf{w}_{b,i} = \lambda^2_{b,u}||\mathbf{w}_{b,i}||_2^2$, where $\lambda^2_{b,u}$ is the long term channel statistics of $\overline{\mathbf{h}}_{b,u}$, $E_{\overline{h}}\left[\overline{\mathbf{h}}^H_{b,u}\overline{\mathbf{h}}_{b,u}\right] = \lambda^2_{b,u}\mathbf{I}_{N_{\text{T}}}$. We assume that the RSSI is reported by the user and the transmit antennas are uncorrelated. For correlated channels, covariance matrices $E_{\overline{h}}\left[\overline{\mathbf{h}}^H_{b,u}\overline{\mathbf{h}}_{b,u}\right] = \mathbf{C}$ can be incorporated in the problem formulation. Finally, the weighted sum

$$\overline{\gamma}_u = \frac{\left|\sum\limits_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2}{E_{\overline{h}}\sum\limits_{i\neq u}\left\{\left|\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}+\sum\limits_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2\right\}+N_0} \tag{3a}$$

$$= \frac{\left|\sum\limits_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2}{\sum\limits_{i\neq u}\left\{\left|\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2+E_{\overline{h}}\left|\sum\limits_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2+2\Re\left\{\left(\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right)^H\sum\limits_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}E_{\overline{h}}\left[\overline{\mathbf{h}}_{b,u}\right]\mathbf{w}_{b,i}\right\}\right\}+N_0} \tag{3b}$$

$$= \frac{\left|\sum\limits_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2}{\sum\limits_{i\neq u}\left\{\left|\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2+E_{\overline{h}}\left|\sum\limits_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2\right\}+N_0} \tag{3c}$$

$$\geq \frac{\left|\sum\limits_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2}{\sum\limits_{i\neq u}\left\{\left|\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2+|\mathcal{B}_i\backslash\mathcal{B}_u|\sum\limits_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}E_{\overline{h}}\left|\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2\right\}+N_0} \tag{3d}$$

$$= \frac{\left|\sum\limits_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2}{\sum\limits_{i\neq u}\left\{\left|\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2+|\mathcal{B}_i\backslash\mathcal{B}_u|\sum\limits_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}\lambda_{b,u}^2||\mathbf{w}_{b,i}||_2^2\right\}+N_0} \triangleq \gamma_u. \tag{3e}$$

rate maximization of $|\mathcal{U}|$ users is evaluated as

$$R_{\text{tot}} = \sum_u \alpha_u \log_2(1+\gamma_u)\ [\text{bps/Hz}], \tag{4}$$

where α_u is a non-negative weight of the uth user.

III. Precoder design

Limited CSI at the central coordination node makes the design of the precoder all the more difficult. In this section, we derive the precoders with limited CSI that also include the long term channel statistics using the SINR definition from (3e).

A. Successive second order cone programming

We propose a SSOCP to solve the problem of precoder design with limited information. SSOCP is based on SCA that allows us to efficiently solve the problem with guaranteed convergence in every iteration. We adopt an optimization framework originally proposed in [14] for linearizing a non-convex constraint that forms a constraint for the useful signal. We also adopt the techniques in [15], [16] for handling the SINR, and reformulate as SOC constraints. The maximization of weighted sum rate R_{tot} with per-antenna power constraint[1] is formulated as

$$\begin{aligned} \underset{\mathbf{w}_{b,u}}{\text{maximize}} \quad & \prod_{u} (1+\gamma_u)^{\alpha_u} \\ \text{subject to} \quad & \sum_{u \in \mathcal{U}_b} |w_{b,u}^{(k)}|^2 \leq P_{\text{max}}, \forall b \in \mathcal{B}_u, k = 1, \ldots, N_{\text{T}}, \end{aligned} \tag{5}$$

where the logarithm being a monotonically non-decreasing function can be removed from the objective, and P_{max} is the maximum transmit power of an antenna of a BS serving a set of $\mathcal{U}_b$ users. This can be recast by letting $t_u = (1+\gamma_u)^{\alpha_u}$ where γ_u is from (3e) and adding a slack variable β_u as

$$\underset{t_u, \beta_u, \mathbf{w}_{b,u}}{\text{maximize}} \quad \prod_{u} t_u \tag{6a}$$

$$\text{subject to} \quad \frac{\left| \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u} \mathbf{w}_{b,u} \right|^2}{\beta_u} \geq t_u^{1/\alpha_u} - 1, \forall u \in \mathcal{U}, \tag{6b}$$

$$\sum_{i \neq u} \left\{ \left| \sum_{b \in \mathcal{B}_i \cap \mathcal{B}_u} \mathbf{h}_{b,u} \mathbf{w}_{b,i} \right|^2 + |\mathcal{B}_i \backslash \mathcal{B}_u| \sum_{b \in \mathcal{B}_i \backslash \mathcal{B}_u} \lambda_{b,u}^2 ||\mathbf{w}_{b,i}||_2^2 \right\} + N_0 \leq \beta_u, \forall u \in \mathcal{U}, \tag{6c}$$

$$\sum_{u \in \mathcal{U}_b} |w_{b,u}^{(k)}|^2 \leq P_{\text{max}}, \forall b \in \mathcal{B}_u, k = 1, \ldots, N_{\text{T}}. \tag{6d}$$

The LHS of (6b) is of the form quadratic over linear, which is convex function, and t_u^{1/α_u} is convex only when $0 < \alpha_u \leq 1$, and concave when $\alpha_u > 1$. Thus, the constraint is non-convex. A concave approximation of the LHS can be obtained as in [14, (6b)], so we define the following expressions

$$p_u \triangleq \Re \left\{ \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u} \mathbf{w}_{b,u} \right\} \quad \text{and} \quad q_u \triangleq \Im \left\{ \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u} \mathbf{w}_{b,u} \right\}. \tag{7}$$

[1]over all data symbols for a given channel realization

By applying the first order Taylor expansion for $\frac{\left(p_u^2+q_u^2\right)}{\beta_u}$ in LHS of (6b) around the local point $\left\{\widetilde{p}_u, \widetilde{q}_u, \widetilde{\beta}_u\right\}, \forall u \in \mathcal{U}$, we get

$$\frac{2\widetilde{p}_u}{\widetilde{\beta}_u}\left(p_u-\widetilde{p}_u\right)+\frac{2\widetilde{q}_u}{\widetilde{\beta}_u}\left(q_u-\widetilde{q}_u\right)+\frac{\widetilde{p}_u^2+\widetilde{q}_u^2}{\widetilde{\beta}_u}\left(1-\left(\frac{\beta_u-\widetilde{\beta}_u}{\widetilde{\beta}_u}\right)\right)+1 \quad \geq \quad t_u^{1/\alpha_u}. \tag{8}$$

When $\alpha_u > 1$, t_u^{1/α_u} in the RHS of (8) is not convex, so it needs to be replaced by its upper bound. Doing as in [14]-[16], with the first order approximation at the point $\widetilde{t}_u$, we obtain

$$t_u^{1/\alpha_u} \leq \widetilde{t}_u^{1/\alpha_u}+\frac{1}{\alpha_u}\widetilde{t}_u^{\frac{1}{\alpha_u}-1}\left(t_u-\widetilde{t}_u\right). \tag{9}$$

Therefore, combining with (8), we get

$$\begin{aligned}\frac{2\widetilde{p}_u}{\widetilde{\beta}_u}\left(p_u-\widetilde{p}_u\right)+\frac{2\widetilde{q}_u}{\widetilde{\beta}_u}\left(q_u-\widetilde{q}_u\right)+\frac{\widetilde{p}_u^2+\widetilde{q}_u^2}{\widetilde{\beta}_u}\left(1-\left(\frac{\beta_u-\widetilde{\beta}_u}{\widetilde{\beta}_u}\right)\right)+1& \\ \geq \widetilde{t}_u^{1/\alpha_u}+\frac{1}{\alpha_u}\widetilde{t}_u^{\frac{1}{\alpha_u}-1}\left(t_u-\widetilde{t}_u\right).&\end{aligned} \tag{10}$$

Now consider (6c) which can be rewritten as an SOC constraint [16]

$$\begin{aligned}\left(\sum_{i\neq u}\left\{\left|\sum_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2+\left|\mathcal{B}_i\backslash\mathcal{B}_u\right|\sum_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}\lambda_{b,u}^2||\mathbf{w}_{b,i}||_2^2\right\}+\left(\sqrt{N_0}\right)^2+\frac{1}{4}\left(\beta_u-1\right)^2\right)^{1/2}& \\ \leq \frac{1}{2}\left(\beta_u+1\right), \forall u \in \mathcal{U}.&\end{aligned} \tag{11}$$

Therefore, the reformulated convex problem for precoder design with the objective of maximizing the geometric mean of t_u becomes

$$\begin{aligned}&\underset{t_u,\beta_u,\mathbf{w}_{b,u}}{\text{maximize}} && \left(\prod_{u=1}^{|\mathcal{U}|}t_u\right)^{1/|\mathcal{U}|} \\ &\text{subject to} && \text{(6d), (10) and (11),}\end{aligned} \tag{12}$$

where the geometric mean is concave, and the exponent does not affect the optimal value. This is performed merely to simplify the implementation. Also, the interfering terms can be collected in a vector as

$$\mathbf{r}_i=\begin{bmatrix}\sum\limits_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\\ \sqrt{|\mathcal{B}_i\backslash\mathcal{B}_u|}\lambda_{b',u}\mathbf{w}_{b',i}\end{bmatrix}, b'\in\mathcal{B}_i\backslash\mathcal{B}_u, \forall i\neq u. \tag{13}$$

The SSOCP with the above simplified notation is summarized in Algorithm 2.

Algorithm 2 SSOCP algorithm for precoder design

To avoid numerical instability, rescale the aggregated channel matrix and the noise power with a factor of the least pathloss such that the SINR is the same.

Set $maxRetries = \text{MAXRETRIES}$, see Fig. 6 for a possible choice.

while $maxRetries$ **do**

  Randomly initialize the non-zero precoding weight, $\mathbf{w}_{b,u}$, from $\mathcal{CN}(0,1)$, and ensure the power of each antenna is limited to $P_{\max}$.

  Calculate γ_u based on (3e), $\forall u$.

  Set $n = 0$

  Evaluate $\widetilde{p}_u^{(n)}$ and $\widetilde{q}_u^{(n)}$ from (7).

  Evaluate $t_u^{(n)} = (1+\gamma_u)^{\alpha_u}$ and $\beta_u^{(n)} = \frac{\left(\widetilde{p}_u^{(n)}\right)^2 + \left(\widetilde{q}_u^{(n)}\right)^2}{t_u^{(n)} - 1}$

  Set $maxIter = \text{MAXITER}$

  while $maxIter$ AND † **do**

    Treat $p_u^{(n)}$ and $q_u^{(n)}$ as *expressions* in CVX [24] which will be used in (10).

    Solve the convex problem (12) as

$$
\begin{aligned}
\underset{t_u, \beta_u, \mathbf{w}_{b,u}}{\textbf{maximize}} \quad & \textbf{geo\_mean}\,(t_u) \\
\text{subject to} \quad & \\
& \left\| \begin{matrix} \mathbf{r}_i \\ \sqrt{N_0} \\ \frac{1}{2}(\beta_u - 1) \end{matrix} \right\|_2 \leq \frac{1}{2}(\beta_u + 1), \\
& \forall i \in \mathcal{U}, \\
& \text{(6d)}, \\
& \text{and}\,(10), \forall u \in \mathcal{U}.
\end{aligned}
$$

    Update: $t_u^{(n+1)} = t_u^{(n)}, \beta_u^{(n+1)} = \beta_u^{(n)}$

    Update: $p_u^{(n+1)} = p_u^{(n)}, q_u^{(n+1)} = q_u^{(n)}$

    Update: $n = n + 1$

    $maxIter = maxIter - 1$

    Evaluate and save the best weighted sum rate achieved so far, as well as the corresponding precoding weights.

  end while

  $maxRetries = maxRetries - 1$

end while

† The weighted sum rate does not improve within a certain tolerance.

The weighted sum rate maximization is a non-convex problem, and the solution may end up as an inefficient local optimum. In order to further improve the solution, we introduce random initialization similar to [23], where we select the best solution out of a number of random initialization. For a given aggregated channel matrix, a small increase in the number of random initializations, as in step 4, increases the probability to find a solution close to the global optimal [23].

The convergence of the proposed SSOCP algorithm closely follows the analysis carried out for the full CSI case in [14]-[16]. Reformulating the SINR constraints to cope with incomplete CSI does not affect the convergence of the SCA. The interference and noise terms are transformed into a SOC from (6c), and the convex function in (6b) is approximated with a linear lower bound at each iteration. This results in an SCA for every iteration, where the objective is monotonically non-decreasing, thereby guaranteeing convergence. In the subsequent section, we apply the branch and bound technique to show that the proposed SSOCP is very close to the optimal.

B. *Optimization via weighted mean square error minimization*

Maximizing the weighted sum rate can be equivalently formulated as minimizing the weighted sum MSE [12], [17]. In this section, we extend this result to the case of limited CSI and efficient backhauling. We derive the precoder based on weighted MSE formulation that accounts for using long term channel statistics in the precoder design when there is limited CSI at the central coordination node. Consider the received signal at the uth user as in (1). The estimated uth user data at the receiver is $\hat{x}_u = a_u y_u$, where $a_u \in \mathbb{C}$ is the receiver weight. The MSE at the uth receiver ξ_u, can be formulated as

$$\begin{aligned}
\xi_u &= E_{x_u,n_u}\left[\left(x_u-\hat{x}_u\right)\left(x_u-\hat{x}_u\right)^H\right] \\
&= E_{x_u,n_u}\left[\left(x_u-a_u y_u\right)\left(x_u-a_u y_u\right)^H\right] \\
&= E_{x_u}\left[x_u x_u^H\right] - a_u^H E_{x_u,n_u}\left[x_u y_u^H\right] - a_u E_{x_u,n_u}\left[y_u x_u^H\right] + a_u a_u^H E_{x_u,n_u}\left[y_u y_u^H\right] \\
&= 1 - a_u^H \sum_{b\in\mathcal{B}_u}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right)^H - a_u \sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u} + a_u a_u^H \widetilde{c}_u. \qquad (14)
\end{aligned}$$

where $E_{x_u}\left[x_u x_u^H\right]=1$ as the user data is zero mean with unit power and $\widetilde{c}_u=E_{x_u,n_u}\left[y_u y_u^H\right]$ is the variance of the received signal which can be evaluated as

$$\begin{aligned}\widetilde{c}_u &= E_{x_u,n_u}\left[y_u y_u^H\right]\\ &= \sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right)^H+\sum_{i\neq u}\sum_{b\in\mathcal{B}_i}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right)^H+E_{n_u}\left[n_u n_u^H\right]\\ &= \sum_{\forall i}\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right)^H+N_0, \end{aligned}\tag{15}$$

where $E_{n_u}\left[n_u n_u^H\right]=N_0$. Similar to (3e), we split the interference terms to incorporate the long term channel statistics for the unknown channel components, and bound the variance to be pessimistic as

$$c_u = N_0+\left|\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2+\sum_{i\neq u}\left\{\left|\sum_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2+|\mathcal{B}_i\backslash\mathcal{B}_u|\sum_{b\in\mathcal{B}_i\backslash\mathcal{B}_u}\lambda_{b,u}^2||\mathbf{w}_{b,i}||_2^2\right\}.\tag{16}$$

The detailed steps are in Appendix A. Therefore, the MSE in (14) has $\widetilde{c}_u$ replaced with c_u. To find the optimal receive weight, $a_u^\star=a_u^{\text{MMSE}}$, we need to take the $\bigtriangledown_{a_u}\left(\xi_u\right)=0$, which works out to be

$$a_u^\star=\sum_{b\in\mathcal{B}_u}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right)^H c_u^{-1}.\tag{17}$$

Therefore, the MMSE is

$$\begin{aligned}\bar{\xi}_u &= 1-\sum_{b\in\mathcal{B}_u}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right)^H c_u^{-1}\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\\ &= 1-a_u^\star\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}=\frac{1}{1+\gamma_u}.\end{aligned}\tag{18}$$

Thus, maximizing the weighted sum rate (4) can be formulated equivalent to a log(MSE) minimization problem [12], [17] as

$$\begin{aligned}&\underset{\mathbf{w}_{b,u}}{\text{minimize}} \quad \sum_u \alpha_u \log_2\bar{\xi}_u\\ &\text{subject to} \quad \text{(6d)}.\end{aligned}\tag{19}$$

where α_u is a non-negative weight. The problem (19) is non-convex, so as a first step to find a tractable local solution we introduce a new variable, $\breve{\xi}_u$, to upper bound the MSE as well as treat the receive scalar, a_u, as an optimization variable. The reformulated problem is

$$\underset{\mathbf{w}_{b,u},a_u,\breve{\xi}_u}{\text{minimize}} \quad \sum_u \alpha_u \log_2\breve{\xi}_u \tag{20a}$$

$$\text{subject to} \quad 1 - a_u^H \sum_{b \in \mathcal{B}_u} (\mathbf{h}_{b,u}\mathbf{w}_{b,u})^H - a_u \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u} + a_u a_u^H c_u \le \breve{\xi}_u, \forall u, \tag{20b}$$

$$\text{(6d)}. \tag{20c}$$

Still the MSE constraint (20b) is not jointly convex with respect to both $\mathbf{w}_{b,u}$ and a_u. However, for a fixed receiver, a_u, the constraint becomes convex. A successive linear approximation of the concave objective is carried out at the point $\widetilde{\xi}_u^{(n)}$ in the nth iteration as

$$\log_2 \breve{\xi}_u^{(n)} \approx \log_2 \widetilde{\xi}_u^{(n)} + d_u^{(n)} \left(\breve{\xi}_u^{(n)} - \widetilde{\xi}_u^{(n)} \right) \log_2 e, \tag{21}$$

where $d_u^{(n)}$ is the linearizing coefficient, which is a non-negative MSE weight for the nth iteration, and it works out to be

$$d_u^{(n)} = \frac{1}{\widetilde{\xi}_u^{(n)}}. \tag{22}$$

Subsituting (21) in the objective (20a), and considering only those terms that depend on $\mathbf{w}_{b,u}$ for a fixed a_u, while ignoring the constant terms and iteration index in (21), the objective becomes

$$\underset{\mathbf{w}_{b,u}, \breve{\xi}_u}{\text{minimize}} \quad \sum_u \alpha_u d_u \breve{\xi}_u. \tag{23}$$

Furthermore, the MSE constraint is tight at the optimal solution. Thus, we replace the MSE upper bound with the actual MSE expression as

$$\zeta = \sum_u \alpha_u d_u \xi_u \tag{24}$$

$$= \sum_u -2\alpha_u \Re \left\{ d_u a_u \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u} \right\} + \sum_u \alpha_u d_u a_u \left\{ \left| \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u} \right|^2 + \sum_{i \neq u} \left\{ \left| \sum_{b \in \mathcal{B}_i \cap \mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,i} \right|^2 + |\mathcal{B}_i \backslash \mathcal{B}_u| \sum_{b \in \mathcal{B}_i \backslash \mathcal{B}_u} \lambda_{b,u}^2 ||\mathbf{w}_{b,i}||_2^2 \right\} \right\}. \tag{25}$$

For fixed a_u, (21) can be solved via successive linearization until convergence. For fixed $\mathbf{w}_{b,u}$, the optimal solution of (21) is given by the optimal receiver weight, a_u in (17). This leads to alternating optimization with monotonic convergence. In practice, we update a_u and then d_u just once without sacrificing monotonicity of the objective. Thus, the original problem (21) can be split as a 3-stage algorithm [12], [17], [18], where the receiver weights, linearizing coefficients, and the precoders are optimized in an alternating manner, i.e., (i) the receiver weights are updated for a given precoder, (ii) the linearizing coefficients are updated for a given precoder, and (iii)

the precoders are evaluated for the given receiver weights and linearizing coefficients. When compared to [12], [17], [18], we design the precoder with limited information and achieve efficient backhauling in a JT-CoMP scenario with per-antenna power constraint. The MSE based precoder design with limited information is outlined in Algorithm 3. The convergence is evaluated based on the MSE of each user as $-\sum_u \alpha_u \log_2 \xi_u$. This is a monotonically non-decreasing function, and the algorithm is terminated when there is no further improvement.

Algorithm 3 The MSE approach for finding the precoder

1: Randomly initialize every precoding weight, $\mathbf{w}_{b,u}$, from $\mathcal{CN}(0,1)$, apply the equivalent limited backhauling, and ensure the power of each antenna is limited to $P_{\max}$.

2: **while** Convergence **do**

3: Evaluate: $a_u, d_u, c_u, \forall u$

4: Solve the convex problem as

$$\begin{aligned} \underset{t_u, \mathbf{w}_{b,u}}{\textbf{minimize}} \quad & \sum_u t_u \\ \text{subject to} \quad & -2\alpha_u \Re\left\{ d_u a_u \sum_{b\in\mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u} \right\} \\ & + \alpha_u d_u a_u a_u^H \left| \sum_{b\in\mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,u} \right|^2 \\ & + \alpha_u d_u a_u a_u^H \left\{ \sum_{i\neq u} \left\{ \left| \sum_{b\in\mathcal{B}_i\cap\mathcal{B}_u} \mathbf{h}_{b,u}\mathbf{w}_{b,i} \right|^2 \right.\right. \\ & \left.\left. + |\mathcal{B}_i \backslash \mathcal{B}_u| \sum_{b\in\mathcal{B}_i\backslash\mathcal{B}_u} \lambda_{b,u}^2 ||\mathbf{w}_{b,i}||_2^2 \right\} \right\} \leq t_u, \forall u \in \mathcal{U}, \\ & \text{(6d).} \end{aligned}$$

Based on precoding weights, steps 5-7 below are applied $\forall u$.

5: Update: receiver variance, c_u, based on (16)

6: Update: receiver weight, a_u, based on (17)

7: Update: linearizing coefficient, d_u, based on (22)

8: **end while**

9: **return** Precoding matrix

The problem of minimizing, ζ, can be solved either with generic solvers, such as those provided with the CVX package [24], where the per-antenna power constraint is formulated as an SOC

program, or solving iteratively via the Karush-Kuhn-Tucker conditions as highlighted in [18, Appx. A]. Note that the latter approach may be preferrable when the number of power constraints (corresponding to the dual variables) is relatively small. The MSE algorithm guarantees convergence as shown in [12, Thm. 3]. The main difference to [12] is that here the receive variance is affected by the long term channel statistics, thus the same convergence analysis applies.

C. Stochastic optimization using particle swarm optimization

The researchers modeling the movement of birds or a shoal of fish discovered that these movements were indeed performing optimization. This gave birth to an entire field of swarm intelligence. In particular, we focus on PSO, a stochastic optimization technique that can provide an acceptable solution even when the problem is non-convex. PSO was proposed to design the precoder in a CoMP setup with limited information [11]. We consider the implementation of PSO as described in [11, Algo. 2], with the addition of random initialization, giving rise to a multi-start PSO such that global optimization can be performed. PSO is a very attractive tool for precoder design as it does not involve any matrix inversion, and the overhead of book keeping the number of variables is very little compared to genetic algorithms. However, being heuristic in nature, it does not guarantee optimality and the algorithm might not converge in polynomial time with the increase in problem size.

D. Branch and Bound

The SSOCP and the MSE reformulated approaches are iterative algorithms where every sub-step is optimal and well justified, leading to monotonic improvement of the objective, with guaranteed convergence. However, these approaches even with a large number of random initializations is not guaranteed to obtain the optimal solution, as the problem is non-convex and NP-hard. Hence, we need to verify that our proposed solution is tightly bounded. In this regard, we consider [13] where branch and bound (BB) is applied for weighted sum rate maximization in MISO downlink cellular networks. We reformulate this problem for joint transmission CoMP. This provides the lower and upper bounds for the problem with full and limited feedback, and also when the pessimistic statistical interference model is used in the precoder design. With this approach, we can say how close the proposed algorithm is from being globally optimum.

When limited CSI information is available at the central coordination node, the branch and bound technique can be applied via reformulating the weighted sum rate maximization in MISO downlink [13] to joint transmission CoMP networks. It is intuitive to observe that the reformulation is exactly the same as [13] however the SINRs are based on (3e). Instead of rewriting the whole branch and bound procedure as described in [13], we highlight the main differences in the reformulated problem. The initialization of the hyperrectangle in [13, (9)] is

$$\mathcal{Q}_{\text{init}} = \left\{ \gamma | 0 \leq \gamma_u \leq |\mathcal{B}_u| \, N_{\text{T}} P_{\max} \sum_{b \in \mathcal{B}_u} ||\mathbf{h}_{b,u}||_2^2 / N_{0, \forall u \in \mathcal{U}} \right\}. \tag{26}$$

The initial hyperrectangle, $Q \in \mathcal{Q}_{\text{init}}$ comprises of $\gamma_{\max} \in \mathbb{R}_0^+$, and $\gamma_{\min} \in \mathbb{R}_0^+, \gamma_{\min} = \mathbf{0}_{|\mathcal{U}|}$. The upper limit in (26) can be obtained for each user, when considering only the SNR from (3e), where the uth user is the only user in the system without any interference. It is intuitive to see that using the Cauchy-Schwartz inequality we have $\left| \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u} \mathbf{w}_{b,u} \right|^2 \leq |\mathcal{B}_u| \sum_{b \in \mathcal{B}_u} |\mathbf{h}_{b,u} \mathbf{w}_{b,u}|^2 \leq |\mathcal{B}_u| \sum_{b \in \mathcal{B}_u} ||\mathbf{h}_{b,u}||_2^2 ||\mathbf{w}_{b,u}||_2^2 \leq |\mathcal{B}_u| \, N_{\text{T}} P_{\max} \sum_{b \in \mathcal{B}_u} ||\mathbf{h}_{b,u}||_2^2$. The last inequality is due to the per-antenna power constraint. The other main contribution lies in the check for the feasibility under limited feedback and backhauling constraint. This is captured in Algorithm 4. The feasibility check is performed as part of the BB technique, and for completeness we provide a cookbook version of BB in Appendix B.

Algorithm 4 Check if $\gamma = \left[\gamma_1, \ldots, \gamma_u, \ldots, \gamma_{|\mathcal{U}|} \right]$ is feasible.

1: Check for feasibility by solving the convex problem

$$\textbf{find} \quad \mathbf{w}_{b,u}, \forall b \in \bigcup_{k \in \mathcal{U}} \mathcal{B}_k, \forall u \in \mathcal{U},$$

subject to

$$\left\| \begin{matrix} \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u} \mathbf{w}_{b,u} \\ \mathbf{r}_i \\ \sqrt{N_0} \end{matrix} \right\|_2 \leq \sqrt{1 + \frac{1}{\gamma_u}} \sum_{b \in \mathcal{B}_u} \mathbf{h}_{b,u} \mathbf{w}_{b,u}, \ddagger$$

$\forall i \in \mathcal{U}$,

(6d).

2: **return** feasibility, and save $\mathbf{w}_{b,u}, \forall b, u$ when feasible

‡ The RHS of this SOCP formulation can be argued along the same lines as [25, Sec. IV.B].

IV. Simulations

We consider $|\mathcal{U}| = 3$ users that are uniformly dropped around the cell-edge at the intersection of $|\mathcal{B}| = 3$ BSs, where each BS has $N_\mathrm{T} = 1, 3$ transmit antenna(s) covering a cell-radius of 500 m. The cell-edge SNR is defined as the SNR experienced by one user at the cell-edge. For simplicity, we set $\alpha_u = 1, \forall u$. The variance of the shadow fading component is $\sigma^2_{\mathrm{SF}} = 8\,\mathrm{dB}$. The receiver noise power is $N_0 = kTB_n$ Watts, where k is the Boltzmann's constant 1.38×10^{-23}Joules/Kelvin, $T = 290$ Kelvin is the operating temperature, and $B_n = 10\,\mathrm{MHz}$ is the system bandwidth.

The legends in the following figures are summarized in Table I. They capture as to how much feedback or backhauling is required or being used based on a given relative threshold. The algorithms without subscripts such as SSOCP, PSO, MSE, ZF, $\mathrm{BB_{UB}}$ and $\mathrm{BB_{LB}}$, capture the case of full feedback and full backhauling, when the relative threshold, $T = \infty\,\mathrm{dB}$. The algorithms with subscripts such as $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$, $\mathrm{MSE}_{\lambda,\mathrm{PL},0}$, $\mathrm{BB}_{\lambda,\mathrm{UB},0}$ and $\mathrm{BB}_{\lambda,\mathrm{LB},0}$ capture the case of limited feedback incorporating our proposed long term channel statistics, and with limited backhauling. The $\mathrm{SSOCP}_{\mathrm{PL},0}$ algorithm captures the naive approach of including the long term channel statistics where they are directly replacing the missing channel coefficients in the interference terms of SINR formulation, when there is limited feedback and limited backhauling. The algorithms with subscripts such as SSOCP_0, PSO_0, $\mathrm{BB}_{\mathrm{UB},0}$ and $\mathrm{BB}_{\mathrm{LB},0}$ capture the case of limited feedback and limited backhauling without the use of any side information.

A. *Effect of threshold and cell-edge SNR*

Fig. 1 shows the *expected* average weighted sum rate evaluated at the central coordination node when designing the precoder for various relative thresholds. It is intuitive to note that with complete information the performance of SSOCP, PSO and ZF are independent of the threshold. With limited information, PSO_0 and SSOCP_0 have similar performance. The most interesting curves are $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$ and $\mathrm{SSOCP}_{\mathrm{PL},0}$, where $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$ incorporates the long term channel statistics in the interference terms, thereby resulting in a pessimistic precoder design. The $\mathrm{SSOCP}_{\mathrm{PL},0}$ naively replaces zeros with these long term channel statistics and appear to achieve superior performance when designing the precoder at the central coordination node. However, it is important to observe the actual performance of the precoder due to the transmission to the user. This is captured in Fig. 2 where $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$ outperforms all other cases when there

Table I

THE INTERPRETATION OF LEGENDS RELATED TO THE INFORMATION AVAILABLE AND GENERATED AT THE CENTRAL COORDINATION NODE IS LISTED BELOW.

Legend†	CSI Feedback	Precoding weights
$\mathrm{BB_{LB}}$	Full	Full
$\mathrm{BB_{UB}}$	Full	Full
$\mathrm{BB}_{\lambda,\mathrm{LB},0}$	Use long term stats (3e)	Limited
$\mathrm{BB}_{\lambda,\mathrm{UB},0}$	Use long term stats (3e)	Limited
$\mathrm{BB_{LB,0}}$	Limited	Limited
$\mathrm{BB_{UB,0}}$	Limited	Limited
MSE	Full	Full
$\mathrm{MSE}_{\lambda,\mathrm{PL},0}$	Use long term stats (25)	Limited
PSO	Full	Full
$\mathrm{PSO_0}$	Limited	Limited
SSOCP	Full	Full
$\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$	Use long term stats (3e)	Limited
$\mathrm{SSOCP_{PL,0}}$	Use long term stats directly	Limited
$\mathrm{SSOCP_0}$	Limited	Limited
ZF	Full	Full

†The acronyms in the legend are summarized here for convenience. The $\mathrm{BB_{UB}}$ and $\mathrm{BB_{LB}}$ denote the upper and lower bound obtained from branch and bound as presented in Algorithm 5, PSO: particle swarm optimization, MSE: weighted mean square error, SSOCP: successive second order cone programming, and ZF: zero forcing.

is limited information. It is interesting to note that the expected rates in Fig. 1 are in line with the actual rates in Fig. 2 for the proposed $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$ approach.

Fig. 3 captures the effect of increasing the cell-edge SNR on the average weighted sum rate for a relative threshold of 9 dB. In the case of having limited information and long term channel statistics, $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$ can be useful compared to $\mathrm{SSOCP_0}$. It is interesting to note that the naive approach $\mathrm{SSOCP_{PL,0}}$ only performs well at low thresholds (not shown here), and the performance deteriorates at high thresholds, and also with the increase in the cell-edge SNR.

Fig. 4 shows the cumulative distribution function (CDF) of the weighted sum rate of the MSE and SSOCP based precoder. The MSE approach achieves a performance similar to that of the SSOCP. The MSE approach is very attractive due to the simple sub-problems being solved in every iteration. However, it takes a longer time for convergence due to the receiver updates.

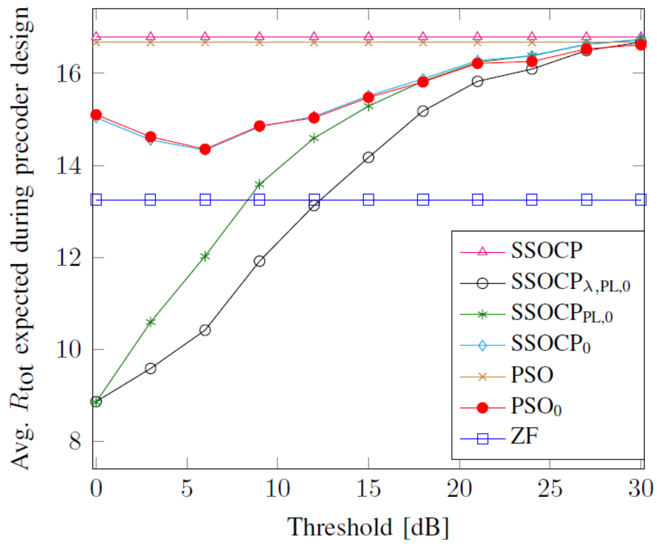

Figure 1. The performance of the precoders in terms of average weighted sum rate, R_{tot} expected when designing the precoder at the central coordination node. The cell-edge SNR is 15 dB, $N_{\text{T}} = 1$, $|\mathcal{B}| = 3$, and $|\mathcal{U}| = 3$.

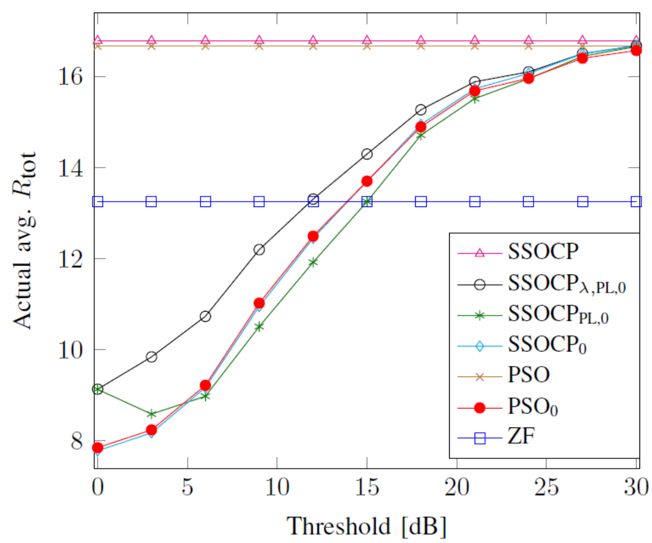

Figure 2. The performance of the precoders in terms of the actual average weighted sum rate, R_{tot} evaluated due to the transmission to the users, with increasing threshold for a given cell-edge SNR of 15 dB, $N_{\text{T}} = 1$, $|\mathcal{B}| = 3$, and $|\mathcal{U}| = 3$.

B. Effect of number of BS antennas

In this section, we investigate the effect of the number of BS antennas on the performance of the precoders. Fig. 5 shows the CDF of the weighted sum rate where the number of antennas is

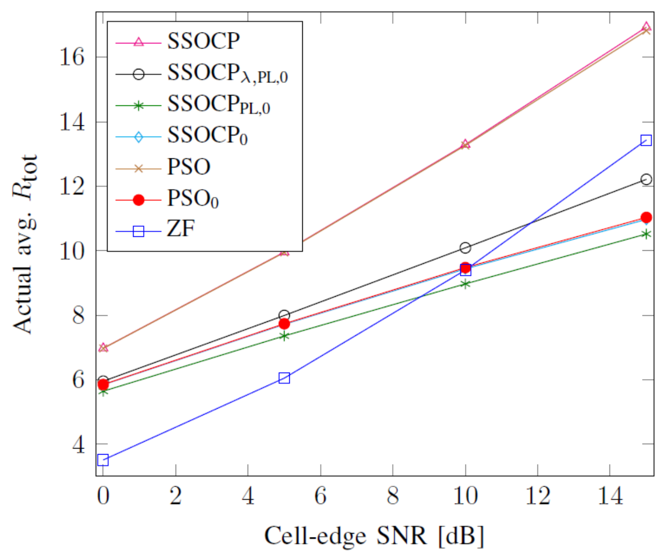

Figure 3. Average weighted sum rate versus cell-edge SNRs for a relative thresholds of 9 dB.

increased to $N_{\mathrm{T}} = 3$ at each of the 3 BSs, serving 9 users, resulting in a fully loaded system. For the non-SSOCP cases, the per-antenna power constraint is applied just as in the case of SSOCP. Apart from which, the whole precoding matrix is rescaled such that at least one of the antennas is transmitting at maximum power. Note that the per-antenna power constraint is more practical and that the SSOCP is more capable of utilizing this to the fullest extent. It can be observed that the proposed SSOCP outperforms all other precoding algorithms. However, the naive approach $\mathrm{SSOCP}_{\mathrm{PL},0}$ performs poorly. It is interesting to note that the SSOCP has consistent cell-edge performance with steeper CDF curves compared to the ZF approach.

Fig. 6 captures the maximum rate that is achieved when designing the precoder at the central coordination node based on the number of random initializations of the precoder. Recall that increasing this number improves the chances of finding a solution close to the global optimum [23]. Each of the subplots in Fig. 6 show the impact on the achievable rate of SSOCP and PSO with the increase in the problem size, due to the increase in the number of antennas at each of the BSs in a fully loaded system. With $N_{\mathrm{T}} = 1$, the 3 BSs serve $|\mathcal{U}| = 3$ users. To keep the system fully loaded, we consider $|\mathcal{U}| = 6$ when $N_{\mathrm{T}} = 2$, and $|\mathcal{U}| = 9$ when $N_{\mathrm{T}} = 3$. For limited feedback, choosing MAXRETRIES $= 5$ is good enough when one considers the tradeoff between the number of initializations and the achievable rate based on the available CSI at the

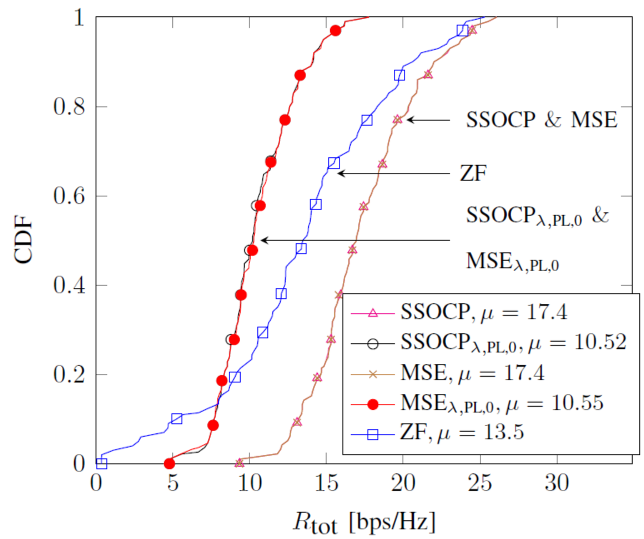

Figure 4. Comparison of the CDF of SSOCP and MSE based precoders in terms of the weighted sum rate. It can be observed that the curves SSOCP and MSE overlap. The SSOCP$_{\lambda,\text{PL},0}$ and MSE$_{\lambda,\text{PL},0}$ curves closely overlap, when long term channel statistics is incorporated under limited CSI and limited backhauling constraint. The cell-edge SNR is 15 dB and the threshold is 3 dB. The μ values in the legend shows the average value.

central coordination node. It is interesting to note that PSO performs closer to SSOCP when the problem size is small, however, SSOCP outperforms consistently with the increase in the problem size in terms of the increase in the number of transmit antennas and the users. It should be noted that the power allocation with PSO is merely a scaling of the entire precoding matrix, as in [8], and that with increased problem size, PSO is unable to completely make use of the per-antenna power constraint as in (6d). With $N_{\text{T}} = 3$, this behavior can be easily explained, and it can be concluded that the PSO is unable to converge in polynomial time. Due to this we do not consider PSO for any further analysis.

C. Bounding the proposed SSOCP

In this section, we show that the accuracy of the solution obtained with SSOCP is tightly bounded. Here, we restrict our simulations to the SSOCP based approach, as we have already observed in Fig. 4 that the MSE and SSOCP have similar performance. Fig. 7 shows the convergence of the branch and bound algorithm under full/limited feedback and backhauling conditions. The y-axis captures the weighted sum rate obtained when designing the precoder at the central coordination node. Notice that the convergence is slow with $(\text{BB}_{\text{UB}}, \text{BB}_{\text{LB}})$ in

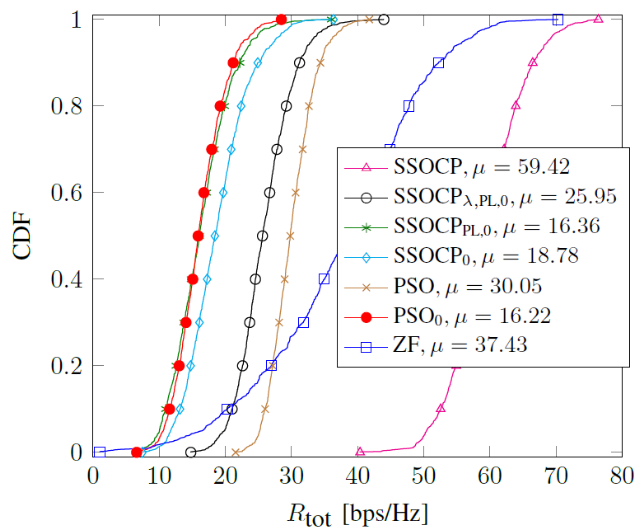

Figure 5. With cell-edge SNR of 15 dB and a threshold of 3 dB, $|\mathcal{B}| = 3$ BSs with $N_\text{T} = 3$ antennas each are serving 9 users. The μ values in the legend shows the average value.

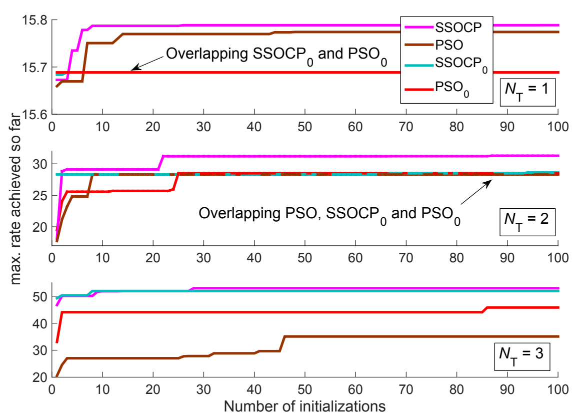

Figure 6. A comparison of the performance of PSO and SSOCP with the number of random initializations for a cell-edge SNR of 15 dB and a relative threshold of 3 dB.

the topmost subplot, when compared to the convergence behavior of the other bounds. This is due to the size of the problem. Also notice that the numerical bounds tightly characterize the proposed SSOCP algorithm for precoder design. It should be mentioned that the bounds can be tightened depending on step 1 in Algorithm 5. The branch and bound technique is extremely slow in the CVX [24] environment even with the bisection method (Algorithm 6) being applied to improve the lower bound. Likewise the CDF of the bounds are captured in Fig. 8 for full

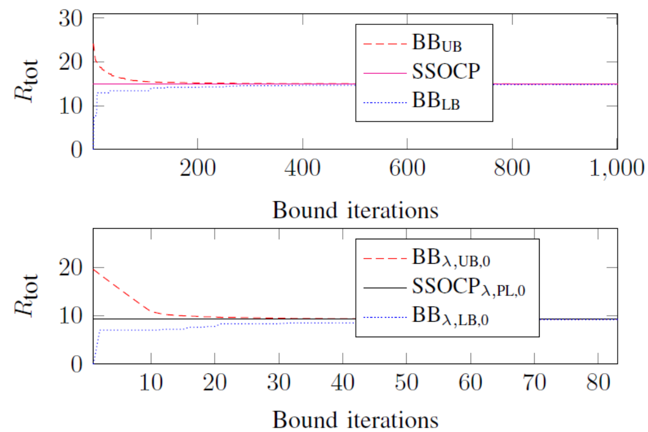

Figure 7. The convergence of the upper and lower bounds from the branch and bound procedure used to benchmark SSOCP and $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$, for a given realization of the aggregated channel matrix with full/limited feedback and backhauling under $|\mathcal{B}| = 3, N_{\mathrm{T}} = 1, |\mathcal{U}| = 3$, with cell-edge SNR of 15 dB and a relative threshold of 3 dB. The MAXRETRIES $= 5, 20$ for SSOCP and $\mathrm{SSOCP}_{\lambda,\mathrm{PL},0}$, respectively.

and limited feedback and backhauling, with and without the use of long term statistics. It can be observed that the BB technique tightly bounds the proposed SSOCP. Note that when limited information is considered, the curves are those that were obtained during the precoder design at the coordination node. It can be observed that when the long term statistics are included, the precoder rate is more pessimistic. However, they perform better during actual transmission when the complete channel is considered as seen in Fig. 2.

V. Conclusions

In this work, we have incorporated the long term channel statistics in the SSOCP algorithm to efficiently solve the precoder design when there is limited channel state information available at the central coordination node. Efficient backhauling is achieved, in a sense that the number of precoding weights generated for the active links is equal to the number of coefficients of the channel state information correspondingly available for the active links at the central coordination node. The above goals are accomplished with the objective of maximizing the weighted sum rate when jointly transmitting to a group of cell-edge users. The efficiency of the solution obtained with the proposed SSOCP algorithm is verified by the tight upper and the lower bounds. The performance of the precoder is also studied for various thresholds, cell-edge SNRs and with

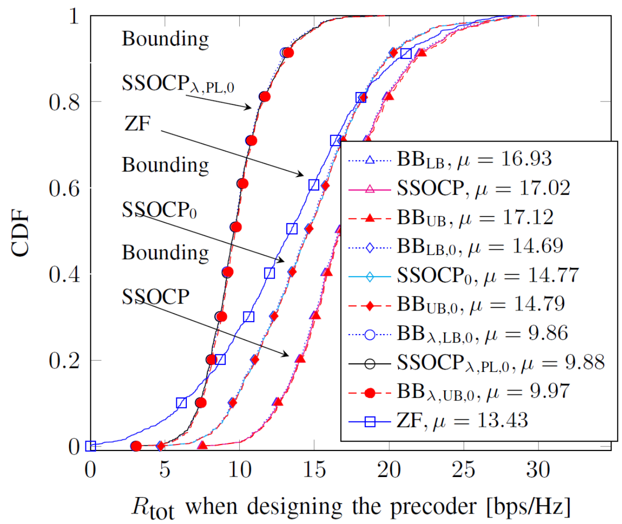

Figure 8. The plot shows the comparison of the CDF of sum rate obtained when designing the precoder and their corresponding bounds. The μ values in the legend shows the average value. Here the MAXRETRIES = 20, in Algorithm 2 for the SSOCP$_{\lambda,\mathrm{PL},0}$, whereas a default value of MAXRETRIES = 5 results in a performance slightly lower than the lower bound.

the increase in the problem size. Alternatively, we also derived the weighted MSE approach for the precoder design that incorporates the long term channel statistics when there is limited information, and it was shown to achieve the performance of SSOCP.

Appendix A
The receive variance with limited CSI and long term channel statistics

To obtain (16), consider (15)

$$\begin{aligned}\widetilde{c}_u &= \sum_{\forall i}\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right)^H + N_0 \\ &= \sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right)^H + \sum_{i\neq u}\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\left(\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right)^H + N_0 \\ &= \left|\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2 + \sum_{i\neq u}\left|\sum_{b\in\mathcal{B}_i}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2 + N_0.\end{aligned}$$

With limited information, we consider the expected value of the inactive links in the interference terms as

$$\begin{aligned}
c_u &= N_0 + \left|\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2 + E_{\overline{h}}\sum_{i\neq u}\left\{\left|\sum_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i} + \sum_{b\in\mathcal{B}_i\setminus\mathcal{B}_u}\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2\right\} \\
&\leq N_0 + \left|\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2 + \sum_{i\neq u}\left\{\left|\sum_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2 + |\mathcal{B}_i\setminus\mathcal{B}_u|\sum_{b\in\mathcal{B}_i\setminus\mathcal{B}_u}E_{\overline{h}}\left|\overline{\mathbf{h}}_{b,u}\mathbf{w}_{b,i}\right|^2\right\} \\
&= N_0 + \left|\sum_{b\in\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,u}\right|^2 + \sum_{i\neq u}\left\{\left|\sum_{b\in\mathcal{B}_i\cap\mathcal{B}_u}\mathbf{h}_{b,u}\mathbf{w}_{b,i}\right|^2 + |\mathcal{B}_i\setminus\mathcal{B}_u|\sum_{b\in\mathcal{B}_i\setminus\mathcal{B}_u}\lambda_{b,u}^2||\mathbf{w}_{b,i}||_2^2\right\}.
\end{aligned}$$

Steps similar to (3a)-(3e) are applied here.

Appendix B

Cookbook version of the branch and bound that incorporates the long term channel statistics

For completeness of Section III-D, a cookbook version of the branch and bound based on [13] is provided in Algorithm 5, with emphasis on including the long term channel statistics into the precoder design, where the SINRs are checked for feasibility in Algorithm 4.

Algorithm 5 The branch and bound algorithm for bounding with limited information, and applying bisection method to improve γ_{max} and return the upper bound BB_{UB} and the lower bound BB_{LB} of the objective in (4).

1: Set tolerance $\epsilon = 0.1$, $maxIter = 100$
2: Set $\mathcal{Q}_{\text{curr}} = \mathcal{Q}_{\text{init}}$
3: Algorithm 6: Bisection method to limit γ_{max} of $\mathcal{Q}_{\text{curr}}$
4: Algorithm 4: Check if γ_{min} of $\mathcal{Q}_{\text{curr}}$ is feasible
5: Algorithm 7: Update BB_{UB}, BB_{LB} based on the above feasibility
6: Accumulate hyperrectangles and the corresponding bounds:
$\mathcal{A} = \{(\mathcal{Q}_{\text{curr}}, B_{\text{UB}}(\mathcal{Q}_{\text{curr}}), B_{\text{LB}}(\mathcal{Q}_{\text{curr}}))\}$
7: **while** $BB_{\text{UB}} - BB_{\text{LB}} > \epsilon$ AND $maxIter$ **do**
8:     **for** $a \in \mathcal{A}$ **do**
9:         **if** $BB_{\text{LB}} = a.B_{\text{LB}}$ **then**
10:             $\mathcal{Q}_{\text{curr}} = a.\mathcal{Q}$
11:             $a_{\text{curr}} = a$
12:             **break**
13:         **end if**
14:     **end for**
    // Branching
15:     Split the longest edge of hyperrectangle $\mathcal{Q}_{\text{curr}}$ into $\mathcal{Q}_1$ and $\mathcal{Q}_2$
16:     **for** $i = 1, 2$ **do**
17:         Algorithm 6: Bisection method to limit γ_{max} of $\mathcal{Q}_i$
18:         Algorithm 4: Check if γ_{min} of $\mathcal{Q}_i$ is feasible
19:         Algorithm 7: Update $B_{\text{UB}}(\mathcal{Q}_i)$ and $B_{\text{LB}}(\mathcal{Q}_i)$
20:     **end for**
21:     Remove $\{a_{\text{curr}}\}$ from $\mathcal{A}$
22:     Update $\mathcal{A} = \mathcal{A} \cup$
    $\{(\mathcal{Q}_1, B_{\text{UB}}(\mathcal{Q}_1), B_{\text{LB}}(\mathcal{Q}_1)), (\mathcal{Q}_2, B_{\text{UB}}(\mathcal{Q}_2), B_{\text{LB}}(\mathcal{Q}_2))\}$
    // Bounding
23:     $BB_{\text{UB}} = \min_{a \in \mathcal{A}} (a.B_{\text{UB}})$
24:     $BB_{\text{LB}} = \min_{a \in \mathcal{A}} (a.B_{\text{LB}})$
25:     $maxIter = maxIter - 1$
26: **end while**
27: **return** $BB_{\text{UB}}, BB_{\text{LB}}, \mathbf{W}^\star$ from step 18

Algorithm 6 Bisection method to improve γ_{max} for the lower bound.

1: Set tolerance $\epsilon = 0.01$
2: **for** Each user **do**
3: $\quad \mathbf{a} = \gamma_{\text{min}} + (\gamma_{u,\text{max}} - \gamma_{u,\text{min}}) \cdot \mathbf{e}_u$ where $\mathbf{e}_u$ is the standard basis vector
4: $\quad$ Set $\mathbf{b}_{\text{lower}} = \gamma_{\text{min}}$ and $\mathbf{b}_{\text{upper}} = \mathbf{a}$
5: $\quad$ Algorithm 4: Check if $\mathbf{a}$ is feasible
6: $\quad$ **if** feasible **then**
7: $\quad\quad \gamma^{\star}_{u,\text{max}} = \mathbf{a}$
8: $\quad\quad$ **continue** step 2 // Avoid unnecessary bisection steps below
9: $\quad$ **end if**
10: $\quad$ **while** $||\mathbf{b}_{\text{upper}} - \mathbf{b}_{\text{lower}}||_2 > \epsilon$ **do**
11: $\quad\quad \mathbf{t} = (\mathbf{b}_{\text{lower}} + \mathbf{b}_{\text{upper}})/2$
12: $\quad\quad$ Algorithm 4: Check if $\mathbf{t}$ is feasible
13: $\quad\quad$ **if** feasible **then**
14: $\quad\quad\quad \mathbf{b}_{\text{lower}} = \mathbf{t}$
15: $\quad\quad$ **else**
16: $\quad\quad\quad \mathbf{b}_{\text{upper}} = \mathbf{t}$
17: $\quad\quad$ **end if**
18: $\quad$ **end while**
19: $\quad \gamma^{\star}_{\mathbf{u},\text{max}} = \mathbf{b}_{\text{upper}}$
20: **end for**
21: **return** $\gamma^{\star}_{\text{max}} = \left[\gamma^{\star}_{1,\text{max}}, \ldots, \gamma^{\star}_{u,\text{max}}, \ldots, \gamma^{\star}_{|\mathcal{U}|,\text{max}}\right]$

Algorithm 7 Update the bounds based on a given hyperrectangle $\mathcal{Q}_{\text{given}}$ and its feasibility.

1: **if** feasible **then**
2: $\quad$ Set $\gamma_u = \gamma_{u,\text{min}}$
3: $\quad$ Evaluate (4), $B_{\text{U}}(\mathcal{Q}_{\text{given}}) = R_{\text{tot}}$
4: $\quad$ Set $\gamma_u = \gamma_{u,\text{max}}$
5: $\quad$ Evaluate (4), $B_{\text{L}}(\mathcal{Q}_{\text{given}}) = R_{\text{tot}}$
6: **else**
7: $\quad$ set $B_{\text{U}}(\mathcal{Q}_{\text{given}}) = 0$ and $B_{\text{L}}(\mathcal{Q}_{\text{given}}) = 0$
8: **end if**
9: **return** $B_{\text{U}}(\mathcal{Q}_{\text{given}}), B_{\text{L}}(\mathcal{Q}_{\text{given}})$

ACKNOWLEDGMENT

We would like to thank the colleagues at the Centre of Wireless Communications, University of Oulu, for the friendly, helpful, and interesting discussions. Also, we would like to thank the members of the VR project meetings with Uppsala University and Karlstad University for their comments. Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. This work is also supported by the Swedish Research Council VR under the project 621-2009-4555 Dynamic Multipoint Wireless Transmission. Some computations were performed on C^3SE.

REFERENCES

[1] S. Shamai (Shitz) and B. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," *IEEE Veh. Technol. Conf.*, 2001.

[2] 3GPP TR 36.819-b20, "3rd Generation Partnership Project, Technical Specification Group Radio Access Network, Coordinated Multi-Point Operation for LTE Physical Layer Aspects (Release 11)," Sep. 2013.

[3] L. Daewon, S. Hanbyul, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated Multipoint Transmission and Reception in LTE-Advanced: Deployment Scenarios and Operational Challenges," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, Feb. 2012.

[4] A. Papadogiannis, H. Bang, D. Gesbert, and E. Hardouin, "Downlink overhead reduction for multicell cooperative processing enabled wireless networks," *IEEE Personal, Indoor and Mobile Radio Commun.*, 2008.

[5] B. Dai and W. Yu, "Sparse beamforming for limited-backhaul network MIMO system via reweighted power minimization," in *Proc. IEEE Globecom*, pp. 1962–1967, Dec. 2013.

[6] A. Papadogiannis, H.J. Bang, D. Gesbert, and E. Hardouin, "Efficient Selective Feedback Design for Multicell Cooperative Networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 196–205, Jan. 2011.

[7] T. Yoo and A. Goldsmith, "On the Optimality of Multiantenna Broadcast Scheduling using Zero-forcing Beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[8] H Zhang and H Dai, "Cochannel Interference Mitigation and Cooperative Processing in Downlink Multicell Multiuser MIMO Networks," *EURASIP J. on Wireless Commun. and Netw.*, vol. 2, pp. 222–235, Dec. 2004.

[9] P. Marsch and G. Fettweis, "A framework for optimizing the downlink of distributed antenna systems under a constrained backhaul," in *Proc. European Wireless Conf.*, Apr. 2007.

[10] J. Zhao, T. Quek, and Z. Lei, "Coordinated Multipoint Transmission with Limited Backhaul Data Transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2627–2775, Jun. 2013.

[11] T. R. Lakshmana, C. Botella, and T. Svensson, "Partial Joint Processing with Efficient Backhauling using Particle Swarm Optimization," *EURASIP J. of Wireless Commun. and Netw.*, vol. 2012, 2012.

[12] Q. Shi, M. Razaviyayn, Z-Q. Luo, and H. Chen, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sept. 2011.

[13] S. Joshi, P. Weeraddana, M. Codreanu, and M. Latva-aho, “Weighted Sum-Rate Maximization for MISO Downlink Cellular Networks via Branch and Bound,” *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2090–2095, Apr. 2012.

[14] G. Venkatraman, A. Tölli, L.N. Tran, and M. Juntti, “Queue aware precoder design for space frequency resource allocation,” *IEEE International Conf. on Acoust., Speech, Signal Process. (ICASSP)*, pp. 860–864, May 2014.

[15] L. N. Tran, M. F. Hanif, A. Tölli, and M. Juntti, “Fast converging algorithm for weighted sum rate maximization in multicell MISO downlink,” *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 872–875, Dec. 2012.

[16] M. F. Hanif, L. N. Tran, A. Tölli, and M. Juntti, “Efficient Solutions for Weighted Sum Rate Maximization in Multicellular Networks with Channel Uncertainties”, *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5659–5674, Nov. 2013.

[17] S.S. Christensen, R. Agarwal, E. Carvalho, and J.M. Cioffi, “Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp.4792–4799, Dec. 2008.

[18] P. Komulainen, A. Tölli, and M. Juntti, “Effective CSI Signaling and Decentralized Beam Coordination in TDD Multi-Cell MIMO Systems,” *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2204–2218, May 2013.

[19] R. Fritzsche and G. Fettweis, ”Robust sum rate maximization in the multi-cell MU-MIMO downlink,” in *Proc. IEEE Wireless Communications and Networking Conf.*, Apr. 2013.

[20] R. Apelfröjd and M. Sternad, “Robust linear precoder for coordinated multipoint joint transmission under limited backhaul with imperfect channel state information,” in *proc. IEEE International Symposium on Wireless Commun. Syst.*, Aug. 2014.

[21] P. Kerret and D. Gesbert, “Sparse precoding in multicell MIMO systems,” in *Proc. IEEE Wireless Communications and Networking Conf.*, Apr. 2012.

[22] K. Donghyun, S. Oh-Soon, S. Illsoo, and L. Kwang, “Channel Feedback Optimization for Network MIMO Systems,” *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3315–3321, Sep. 2012.

[23] A. Tolli, M. Codreanu, and M. Juntti, “Cooperative MIMO-OFDM Cellular System with Soft Handover Between Distributed Base Station Antennas,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1428–1440, Apr. 2008.

[24] M. Grant and Stephen Boyd, “CVX: Matlab software for disciplined convex programming,” version 2.0 beta, Sep. 2013. [Online] Available: http://cvxr.com/cvx, [Accessed: Aug. 2014].

[25] A. Wiesel, Y. C. Eldar, and S. Shamai, “Linear precoding via conic optimization for fixed MIMO receivers,” *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.